Topic 22: Regression with a Single Regressor: Hypothesis Tests and Confidence Intervals

Test ID: 8450964

Question #1 of 17

Question ID: 438947

Consider the regression results from the regression of Y against X for 50 observations:

> Y = 0.78 + 1.2 X
>
> The standard error of the estimate is 0.40 and the standard error of the coefficient is 0.45.

Which of the following reports the correct value of the *t*-statistic for the slope and correctly evaluates its statistical significance with 95 percent confidence?

- ✗ **A)** ***t* = 1.789; slope is not significantly different from zero.**
- ✗ **B)** *t* = 1.200; slope not significantly different from zero.
- ✗ **C)** *t* = 3.000; slope is significantly different from zero.
- ✓ **D)** *t* = 2.667; slope is significantly different from zero.

Explanation

The test statistic is *t* = (1.2 - 0) / 0.45 = 2.667. The critical *t*-values for 48 degrees of freedom are ± 2.011. Therefore, the slope is different from zero.

Question #2 of 17

Question ID: 438959

If the variance of the residuals is not constant across all observations in the sample, the regression exhibits heteroskedasticity. Effects of heteroskedasticity include which of the following problems?

I. The coefficient estimates in the regression model are affected.

II. If the standard errors are too small, but the coefficient estimates themselves are not affected, the t-statistics may be too large and the null hypothesis of no statistical significance will be rejected too often.

- ✓ **A)** **II only.**
- ✗ **B)** I only.
- ✗ **C)** Both I and II.
- ✗ **D)** Neither I nor II.

Explanation

Effects of heteroskedasticity include:

- The standard errors are usually unreliable estimates.
- The coefficient estimates are not affected.
- If the standard errors are too small, but the coefficient estimates themselves are not affected, the t-statistics will be too large and the null hypothesis of no statistical significance is rejected too often. The opposite will be true if the standard errors are too large.

Question #3 of 17

Question ID: 438946

An analyst is regressing fund returns against the return on the Wilshire 5000 to determine whether beta is equal to 1.0. The analyst is trying to determine whether the number of observations should be increased. Which of the following is a reason why the test will have *higher* power if the number of observations is *increased*? The:

- ✗ **A)** **mean squared error of the regression will be lower.**
- ✗ **B)** constant of the regression will be closer to zero.
- ✗ **C)** estimate of beta will be farther away from 1.0.
- ✓ **D)** standard error of the regression will be lower.

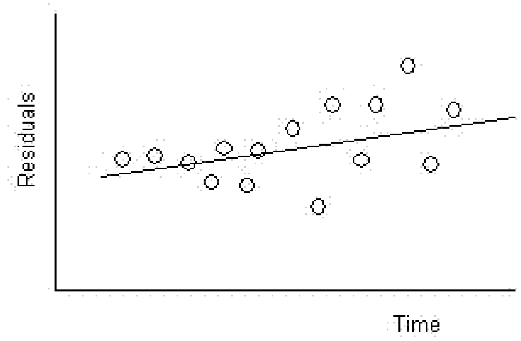
Explanation

A larger number of observations will decrease the standard error of the regression which will increase the size of the test statistic if beta is different than 1.0.

Question #4 of 17

Question ID: 438954

Consider the following graph of residuals and the regression line from a time-series regression:



These residuals exhibit the regression problem of:

- ✗ **A)** **homoskedasticity.**
- ✗ **B)** multicolinearity.
- ✗ **C)** autocorrelation.
- ✓ **D)** heteroskedasticity.

Explanation

The residuals appear to be from two different distributions over time; in the earlier periods, the model fits rather well compared to the later periods.

Question #5 of 17

Question ID: 438958

Which expression best represents the condition homoskedasticity? (In the expressions assume $\sigma^2 > 0$)

✗ **A)** **corr(X_i, ε_i) = 0.**

✗ **B)** $E(\varepsilon_i|X_i) = \sigma^2$.

✗ **C)** corr(ε_i, ε_{i+j}) = 0.

✓ **D)** $V(\varepsilon_i|X_i) = \sigma^2$.

Explanation

Homoskedasticity means the variance of ε_i is constant and unrelated to the value of the independent variable.

Question #6 of 17

Question ID: 438945

A sample of 200 monthly observations is used to run a simple linear regression: Returns = $b_0 + b_1$Leverage + u. The t-value for the regression coefficient of leverage is calculated as t = - 1.09. A 5 percent level of significance is used to test whether leverage has a significant influence on returns. The correct decision is to:

✗ **A)** **reject the null hypothesis and conclude that leverage significantly explains returns.**

✗ **B)** do not reject the null hypothesis and conclude that leverage significantly explains returns.

✗ **C)** reject the null hypothesis and conclude that leverage does not significantly explain returns.

✓ **D)** do not reject the null hypothesis and conclude that leverage does not significantly explain returns.

Explanation

Do not reject the null since |-1.09|<1.96(critical t-value).

Question #7 of 17

Question ID: 438955

An analyst is estimating whether a fund's excess return for a quarter is related to interest rates and last quarter's excess return. The regression equation is found to have unconditional heteroskedasticity and serial correlation. Which of the following is *most* accurate? Parameter estimates will be:

✗ **A)** **accurate but statistical inference about the parameters will not be valid.**

✓ **B)** inaccurate and statistical inference about the parameters will not be valid.

✗ **C)** inaccurate but statistical inference about the parameters will be valid.

✗ **D)** accurate and statistical inference about the parameters will be valid.

Explanation

One of the independent variables is a lagged value of the dependent variable. This means that serial correlation will cause an inaccurate parameter estimate. Serial correlation always impacts the statistical inference about the parameters. Unconditional heteroskedasticity never impacts statistical inference or parameter accuracy.

Question #8 of 17

Question ID: 438944

Consider the following estimated regression equation:

> $ROE_t = 0.23 - 1.50\ CE_t$

The standard error of the slope coefficient is 0.40 and the number of observations is 32. The 95 percent confidence interval for the slope coefficient, b_1, is:

- ✗ **A)** **{$-2.300 < b_1 < -0.700$}.**
- ✗ **B)** {$0.683 < b_1 < 2.317$}.
- ✗ **C)** {$-3.542 < b_1 < 0.542$}.
- ✓ **D)** {$-2.317 < b_1 < -0.683$}.

Explanation

The confidence interval is -1.50 ± 2.042 (0.40), or {$-2.317 < b_1 < -0.683$}.

Questions #9-10 of 17

Assume you ran a multiple regression to gain a better understanding of the relationship between lumber sales, housing starts, and commercial construction. The regression uses lumber sales as the dependent variable with housing starts and commercial construction as the independent variables. The results of the regression are:

	Coefficient	*Standard Error*	*t-statistics*
Intercept	5.37	1.71	3.14
Housing starts	0.76	0.09	8.44
Commercial construction	1.25	0.33	3.78

The level of significance for a 95% confidence level is 1.96

Question #9 of 17

Question ID: 438951

Construct a 95% confidence interval for the slope coefficient for Housing Starts.

- ✓ **A)** **0.76 ± 1.96(0.09).**
- ✗ **B)** 1.25 ± 1.96(0.33).
- ✗ **C)** 0.76 ± 1.96(8.44).
- ✗ **D)** 1.25 ± 1.96(3.78).

Explanation

The confidence interval for the slope coefficient is $b_1 \pm (t_c \times s_{b_1})$.

Question #10 of 17

Question ID: 438952

Construct a 95% confidence interval for the slope coefficient for Commercial Construction.

- ✗ **A)** **0.76 ± 1.96(8.44).**
- ✓ **B)** 1.25 ± 1.96(0.33).
- ✗ **C)** 0.76 ± 1.96(0.09).
- ✗ **D)** 1.25 ± 1.96(3.78).

Explanation

The confidence interval for the slope coefficient is $b_1 \pm (t_c \times s_{b_1})$.

Question #11 of 17

Question ID: 438953

An analyst has been assigned the task of evaluating revenue growth for an online education provider company that specializes in training adult students. She has gathered information about student ages, number of courses offered to all students each year, years of experience, annual income and type of college degrees, if any. A regression of annual dollar revenue on the number of courses offered each year yields the results shown below.

Coefficient Estimates		
Predictor	Coefficient	Standard Error of the Coefficient
Intercept	0.10	0.50
Slope (Number of Courses)	2.20	0.60

Which statement about the slope coefficient is *most correct*, assuming a 5 percent level of significance and 50 observations?

- ✗ **A) t-Statistic: 0.20. Slope: Significantly different from zero.**
- ✓ **B)** t-Statistic: 3.67. Slope: Significantly different from zero.
- ✗ **C)** t-Statistic: 3.67. Slope: Not significantly different from zero.
- ✗ **D)** t-Statistic: 0.20. Slope: Not significantly different from zero.

Explanation

$t = 2.20/0.60 = 3.67$. Since the t-statistic is larger than an assumed critical value of about 2.0, the slope coefficient is statistically significant.

Question #12 of 17

Question ID: 438949

Consider the following estimated regression equation:

> $AUTO_t = 0.89 + 1.32\ PI_t$

The standard error of the coefficient is 0.42 and the number of observations is 22. The 95 percent confidence interval for the slope coefficient, b_1, is:

- ✗ **A) $\{-0.766 < b_1 < 3.406\}$.**
- ✗ **B)** $\{0.900 < b_1 < 1.740\}$.
- ✗ **C)** $\{0.480 < b_1 < 2.160\}$.
- ✓ **D)** $\{0.444 < b_1 < 2.196\}$.

Explanation

The degrees of freedom are found by n-k-1 with k being the number of independent variables or 1 in this case. DF = 22-1-1 = 20. Looking up 20 degrees of freedom on the student's t distribution for a 95% confidence level and a 2 tailed test gives us a critical value of 2.086. The confidence interval is $1.32 \pm 2.086\ (0.42)$, or $\{0.444 < b_1 < 2.196\}$.

Question #13 of 17

Question ID: 438956

Which of the following conditions will *least likely* affect the statistical inference about regression parameters by itself?

- ✗ **A)** **Serial correlation.**
- ✗ **B)** Conditional heteroskedasticity.
- ✗ **C)** Multicollinearity.
- ✓ **D)** Unconditional heteroskedasticity.

Explanation

Unconditional heteroskedasticity does not impact the statistical inference concerning the parameters.

Question #14 of 17

Question ID: 438957

Which of the following statements regarding heteroskedasticity is **FALSE**?

- ✗ **A)** **Conditional heteroskedasticity is the case in which the residuals are correlated with the values of the independent variables.**
- ✓ **B)** The assumption of linear regression is that the residuals are heteroskedastic.
- ✗ **C)** Heteroskedasticity may occur in cross-section or time-series analyses.
- ✗ **D)** Heteroskedasticity results in an estimated variance that is too large and, therefore, affects statistical inference.

Explanation

The assumption of regression is that the residuals are homoskedastic (i.e., the residuals are drawn from the same distribution).

Question #15 of 17

Question ID: 438948

Consider the regression results from the regression of Y against X for 50 observations:

> Y = 0.78 - 1.5 X
>
> The standard error of the estimate is 0.40 and the standard error of the coefficient is 0.45.

Which of the following reports the correct value of the *t*-statistic for the slope and correctly evaluates H_0: $b_1 \geq 0$ versus H_a: $b_1 < 0$ with 95 percent confidence?

- ✓ **A)** ***t* = -3.333; slope is significantly negative.**
- ✗ **B)** *t* = -3.750; slope is significantly different from zero.
- ✗ **C)** *t* = 3.750; slope is significantly different from zero.
- ✗ **D)** *t* = 3.333; slope not significantly different from zero.

Explanation

The test statistic is *t* = (-1.5 - 0) / 0.45 = -3.333. The critical *t*-value for 48 degrees of freedom is +/- 1.667. However, in the Schweser Notes you should use the closest degrees of freedom number of 40 df. which is +/-1.684. Therefore, the slope is different from zero. We reject the null in favor of the alternative.

Question #16 of 17

Question ID: 438960

The Gauss-Markov theorem says that if the linear regression model assumptions are true and the regression errors display homoskedasticity, then the ordinary least squares (OLS) estimators exhibit which of the following properties?

- ✓ **A)** **In repeated sampling, the averages of the coefficients from a sample will be distributed around the true population parameters.**
- ✗ **B)** The OLS estimated coefficients are based on non-linear functions.
- ✗ **C)** The OLS estimated coefficients have the maximum variance compared to other methods of estimating the coefficients.
- ✗ **D)** The OLS estimate of the variance of the errors is biased.

Explanation

If the estimator is the best available (has the minimum variance), exhibits linearity, and is unbiased, it is said to be the best linear unbiased estimator. The OLS estimated coefficients are unbiased, which means that in repeated sampling the averages of the coefficients from the sample will be distributed around the true population parameters.

Question #17 of 17

Question ID: 438943

The estimated slope coefficient from a single linear regression model is 0.55 with a standard error of 0.30. Assuming the sample for this model has 1,000 observations, what can we conclude about the 95% confidence interval for the model's slope coefficient?

- ✗ **A)** **The null hypothesis that the slope coefficient is different than zero should be accepted.**
- ✗ **B)** The slope coefficient is significantly different from zero.
- ✓ **C)** The slope coefficient is not significantly different from zero.
- ✗ **D)** The null hypothesis that the slope coefficient is equal to zero should be accepted.

Explanation

The confidence interval for b_1 is:

$b_1 \pm (t_c \times s_{b1})$

The critical two-tail t-values are roughly ± 1.96 (with a high number of observations). The 95% confidence interval is computed as:

$0.55 \pm (1.96 \times 0.3) = 0.55 \pm 0.588$ or -0.038 to 1.138

Because this confidence interval includes zero, we can conclude that the slope coefficient is not significantly different from zero. In hypothesis testing, we either reject or fail to reject the null hypothesis.